

## REPORT

# *ELMOD2* Is a Candidate Gene for Familial Idiopathic Pulmonary Fibrosis

Ulla Hodgson, Ville Pulkkinen, Morag Dixon, Myriam Peyrard-Janvid, Marko Rehn, Päivi Lahermo, Vesa Ollikainen, Kaisa Salmenkivi, Vuokko Kinnula, Juha Kere, Pentti Tukiainen, and Tarja Laitinen

We performed a genomewide scan in six multiplex families with familial idiopathic pulmonary fibrosis (IPF) who originated from southeastern Finland. The majority of the Finnish multiplex families were clustered in the region, and the population history suggested that the clustering might be explained by an ancestor shared among the patients. The genomewide scan identified five loci of interest. The hierarchical fine mapping in an extended data set with 24 families originating from the same geographic region revealed a shared 110 kb to 13 Mb haplotype on chromosome 4q31, which was significantly more frequent among the patients than in population-based controls (odds ratio 6.3; 95% CI 2.5–15.9;  $P = .0001$ ). The shared haplotype harbored two functionally uncharacterized genes, *ELMOD2* and *LOC152586*, of which only *ELMOD2* was expressed in lung and showed significantly decreased messenger-RNA expression in IPF lung ( $n = 6$ ) when compared with that of healthy lung ( $n = 7$ ;  $P = .05$ ). Our results suggest *ELMOD2* as a novel candidate gene for susceptibility in familial IPF.

Idiopathic pulmonary fibrosis (IPF [MIM 178500]) is a chronic, late-onset disease of lung parenchyma with unknown etiology. IPF has been treated with corticosteroids and immunosuppressive agents, but the prognosis and the response to treatment have remained poor,<sup>1</sup> and the estimated time of survival from the diagnosis is <3 years.<sup>2</sup>

The genetic component in the development of the disease is poorly understood, but, since the 1950s, many anecdotal cases of familial IPF (FPF) have been described in the literature.<sup>3–8</sup> According to our nationwide study, FPF explained 3%–4% of all IPF cases in Finland, which is slightly more than in the United Kingdom (0.5%–2.2%).<sup>6,9</sup> In both countries, the overall prevalence of IPF was at the same level (in Finland, 16–18 in 100,000).<sup>9</sup> Although the mode of inheritance of FPF is unknown, an autosomal dominant model with reduced penetrance has been proposed.<sup>6,10</sup> The structures of the multiplex families that we identified in Finland support this hypothesis.<sup>9</sup>

In Finland, both sporadic and familial cases were found more frequently in southeastern Finland (Savo Province) than anywhere else in the country. Of study families, 65% originated from three rural municipalities where the prevalence of FPF was 50-fold higher than in the rest of the country. On the basis of Finnish population history, we know that the villages in this area are <50 generations old and remained relatively isolated until the 20th century. The isolation resulted in local enrichments of especially recessive but also dominant mutations that cause late-onset diseases.<sup>11</sup> The strong clustering led us to hypothesize that at least some of the patients with IPF may share a common ancestor. This allowed us to further assume that

the patients share not only the same mutation but also the same ancestral chromosomal region of up to 10 cM around the mutation.<sup>12,13</sup> In this situation, a genomewide linkage analysis combined with haplotype association analysis within the linkage peak should be well able to locate a susceptibility gene for IPF.

To map the loci linked to IPF, we performed a genomewide search in six pedigrees. Five of these pedigrees originated from the cluster of FPF.<sup>9</sup> Three loci with the highest NPL and two loci with a potential shared haplotype were studied further with a dense set of markers, and the family cohort was expanded to 24 families originating from the same or nearby areas. To test for the best observed haplotype association, the family data set was further expanded to 35 families, including families from across the country.

For the genomewide search to localize susceptibility genes, we used patients with IPF whose diagnoses had been verified according to the American Thoracic Society/European Respiratory Society diagnostic criteria.<sup>9,14</sup> Their clinical characteristics are shown in table A1 (online only). For the other family members who were genotyped, the physician-diagnosed IPF was excluded by an interview (discussed in detail in the online-only appendix). All participants signed an informed consent form and donated their blood samples. The study was approved by the Ethics Committee of Helsinki University Hospital, Helsinki, and by the Finnish Ministry of Social Affairs and Health.

In the nationwide epidemiological study of IPF, we identified 17 multiplex families in Finland.<sup>9</sup> In six of those families, there were two or three affected siblings. All these

From the Department of Pulmonary Medicine (U.H.; V.K.; P.T.) and HUSLAB (K.S.), Helsinki University Central Hospital, Departments of Medical Genetics (U.H.; V.P.; M.D.; J.K.; T.L.), Computer Science (V.O.), and Pathology (K.S.), and Finnish Genome Center (P.L.; V.O.), University of Helsinki, and GeneOS Oy (M.R.), Helsinki; and Department of Biosciences at Novum, Karolinska Institutet, Stockholm (M.P.-J.; J.K.)

Received December 29, 2005; accepted for publication March 27, 2006; electronically published May 9, 2006.

Address for correspondence and reprints: Dr. Ulla Hodgson, Helsinki University Central Hospital, Haartmaninkatu 4, PB 340, 00029 Helsinki, Finland. E-mail: [ulla.hodgson@hus.fi](mailto:ulla.hodgson@hus.fi)

*Am. J. Hum. Genet.* 2006;79:149–154. © 2006 by The American Society of Human Genetics. All rights reserved. 0002-9297/2006/7901-0017\$15.00

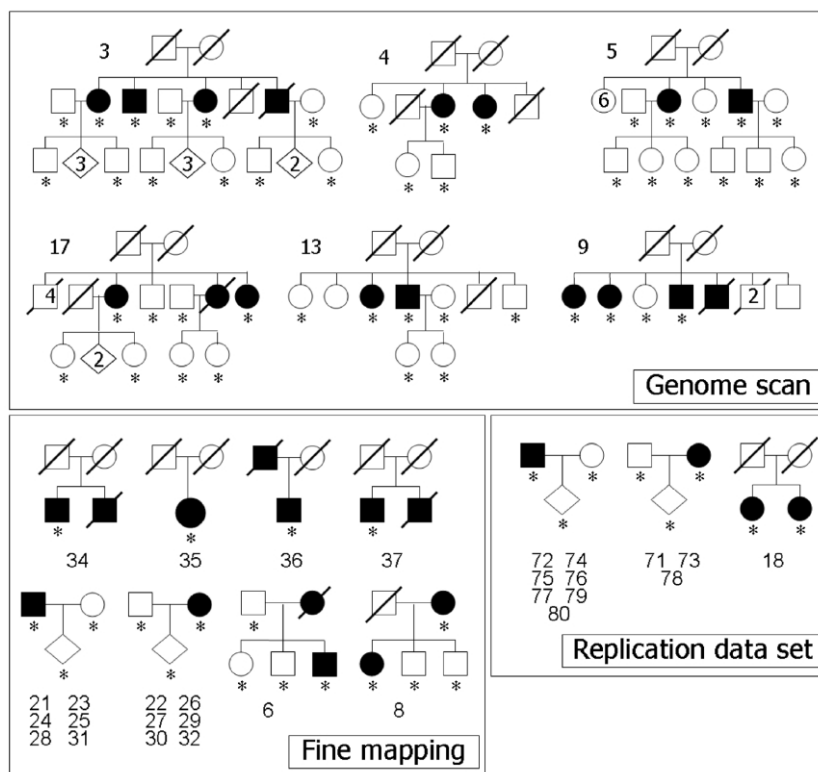
families were included in the genomewide scan without further selection (fig. 1). Genotyping was performed for 14 patients and their 33 first-degree relatives. Five of the six pedigrees, traced for 3–5 generations, originated from the three municipalities in Savo where we previously observed a 50-fold increased risk of FPF compared with the rest of the country.<sup>9</sup> The remaining pedigree originated 150 km west of the geographic cluster of affected population. On the basis of genealogy, no links between the families were found within the last 3–5 generations.

The genomewide scan was performed, with a commercially available microsatellite marker set (Applied Biosystems Linkage Mapping Set MD-10), with an average interval between the markers of 10.6 cM. The genotyping success rate was 83%. NPL analysis was performed using GENEHUNTER 2.0.<sup>15</sup> None of the markers or families showed an exceptional tendency for Mendelian errors. Three loci (fig. A1 [online only]) that exceeded NPL 1.5 were localized on chromosome 3 at marker *D3S1278* (NPL 1.7;  $P = .06$ ; information content 57%), on chromosome 4 at marker *D4S424* (NPL 1.7;  $P = .05$ ; information 60%), and on chromosome 13 at marker *D13S265* (NPL 1.6;  $P = .06$ ; information 71%).

Genomewide mapping in IPF is challenging for several reasons. Because of the short life expectancy of the pa-

tients, the enrollment period for recruiting informative families for linkage analyses is very short. Since IPF is a late-onset disease, all family members aged <60–70 years will have uncertain affection status. IPF is a rare disease, and it is most likely that a portion of familial cases go unnoticed because of the insufficient or noninformative pedigree history and changes in clinical practice and diagnostics over the years. The small number of the study pedigrees with limited information for linkage suggested that it is still possible to miss a true disease locus. Therefore, we visually inspected the haplotypes reconstructed by GENEHUNTER in all the chromosomal regions that showed positive NPL scores and were potentially shared by the affected individuals within a family and across families. Two chromosomal regions appeared to be of interest. On chromosomes 9 and 12, several patients shared the same haplotype, up to 30 cM in length, identified by 1–4 consecutive markers.

To maximize the information content for linkage and to detect or exclude a potential haplotype association, we used the principle of hierarchical genotyping in the extended family data set (fig. 1, “Genome scan” and “Fine mapping” panels). Five loci on chromosomes 3, 4, 9, 12, and 13 were selected for fine mapping. For fine mapping, we genotyped 63 additional microsatellite markers and



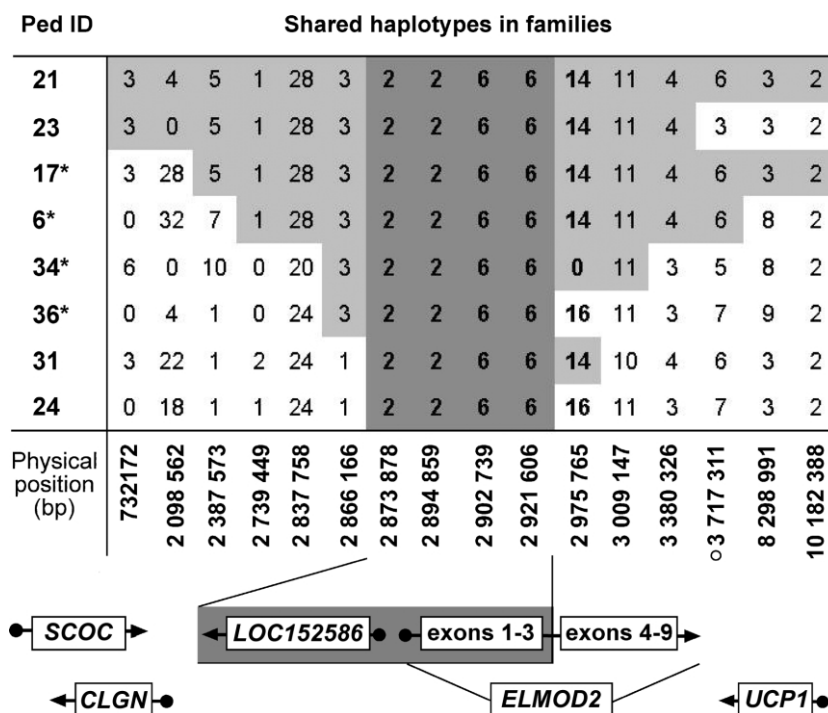
**Figure 1.** Pedigree structures in the genomewide scan ( $n = 6$  pedigrees), fine mapping (Genome scan + Fine mapping;  $n = 24$  pedigrees), and susceptibility-haplotype association study on chromosome 4 (Genome scan + Fine mapping + Replication data set;  $n = 35$  pedigrees). Affected family members are indicated by blackened symbols. Individuals who donated their blood sample for genotyping are marked with an asterisk (\*).

one SNP; the order and average intervals between the markers were estimated from the Human Genome Sequence. All the multiplex families that were not suited for the linkage analysis were included, as were nuclear families who originated from the same region where the enrichment of the FPF was detected but in whom we were not able to confirm the familial background of the disease. A total of 33 patients and their 60 unaffected family members were genotyped. We added a total of 9 markers to 3q13; 28 markers to 4q31, 6 markers to 9q21, 4 markers to 12p12-q12, and 17 markers to 13q31 (table A2 [online only]). With the added markers, the linkage became stronger on chromosomes 3 (at *D3S1303*, NPL 2.07;  $P = .05$ ; information 83%), chromosome 4 (at *D4S1586*, NPL 2.09;  $P = .02$ ; information 74%), and chromosome 13 (at *D13S281*, NPL 2.4;  $P = .01$ ; information 84%), but it weakened on chromosome 12 (at *D12S310*, NPL 0.1;  $P = .4$ ; information 67%) and on chromosome 9 (at *D9S167*, NPL 1.0;  $P = .2$ ; information 79%).

We then considered haplotypes shared between patients. In contrast to any other loci, on chromosome 4, the patients shared a haplotype in 8 of the 24 families (fig. 2). The shared haplotype between the families was 110 kb (fig. 2, dark gray box) at its shortest and was 13 cM (fig.

2, light gray box) at its longest, defined by 4–16 consecutive highly informative markers. The susceptibility haplotype, with shared alleles 2\*2\*6\*6\*, was delimited, both upstream and downstream, by two recombinations. In addition to the weakened linkage results on chromosomes 9 and 12, the putative shared haplotypes were broken down and the chromosomes excluded. On chromosome 3 with 9 additional markers (marker density, on average, 4.4 cM), no evidence of a shared haplotype between patients was found, and it was excluded. On chromosome 13, a conserved haplotype among patients was observed, but the same haplotype was as common among the unaffected family members so showed no evidence of a disease association.

We estimated the frequency of susceptibility haplotype 2\*2\*6\*6\* with markers 2902739 and 2921606 (named according to their positions on genomic clone NT\_016606.15 [Entrez Nucleotide]) for alleles 6\*6\* (fig. 2), which were always inherited in linkage disequilibrium and were sufficiently informative to tag haplotype 2\*2\*6\*6\*. The data set was expanded, with all the families recognized during the study as a small replication data set (one affected sib pair and 10 uniplex families from across Finland). In total, we genotyped 45 patients from 35 families (13 multiplex



**Figure 2.** The shared susceptibility haplotype identified, in 8 of 24 families, within the linkage peak on chromosome 4 (upper panel) and the genes located within or in the immediate vicinity of the susceptibility haplotype, according to the Genome Browser (lower panel). The genotyped alleles are shown as positions on genomic clone NT\_016606.15. The shared haplotype at its shortest is shown with dark gray and is shown at its longest with light gray. The marker that showed the best linkage is marked with a circle (○). Multiplex families are marked with an asterisk (\*). *LOC152586* and the first three exons of *ELMOD2* are located within the haplotype (shown in gray). Short coiled-coil protein (*SCOC*), calmeglin (*CLGN* [MIM 601858]), and uncoupling protein 1 (*UCP1* [MIM 113730]) are located outside the shared haplotype.

and 22 uniplex) (fig. 1). As controls, we used 50 regional individuals (including the spouses of the probands) and 93 blood donors from across Finland. In 38% (5 of 13) of multiplex families, the affected family members shared the susceptibility haplotype (fig. 1, families 6, 17, 18, 34, and 36). Four of these families originated from the cluster of three municipalities. Correspondingly, in 27% (7 of 22) of uniplex families (fig. 1, families 21, 23, 24, 31, 73, 74, and 77), the proband was a haplotype 2\*2\*6\*6\* carrier, and four of these families originated from the cluster. Among the control individuals from Savo Province, the heterozygous carriership of the haplotype was 4% (2 of 50) among controls from the region, and 9.6% (9 of 93) among the controls from across Finland. The carriership of the susceptibility haplotype was significantly higher among the patients (12 of 35) when compared with the regional controls (2 of 50 individuals; odds ratio [OR] 12.5; 95% CI 2.6–60.6;  $P = .0004$ ) and with the combined pool of controls (11 of 143 individuals; OR 6.3; 95% CI 2.5–15.9;  $P = .0001$ ).

To fully explore the genetic variation in the associated 110-kb haplotype shared by several families, we sequenced the nonrepetitive DNA segments in this interval (from position 2866166 to position 2975765 in NT\_016606.15) in two individuals (an affected father and his daughter), both heterozygous carriers of the haplotype. The chromatograms were inspected visually by two independent readers, and sequences were aligned using the Pregap and gap4 software from the Staden package. These sequences were compared with NT\_016606.15. We observed a total of 37 SNPs and two small insertion/deletion polymorphisms (table A3 [online only]). Most of the polymorphisms have not been reported elsewhere.

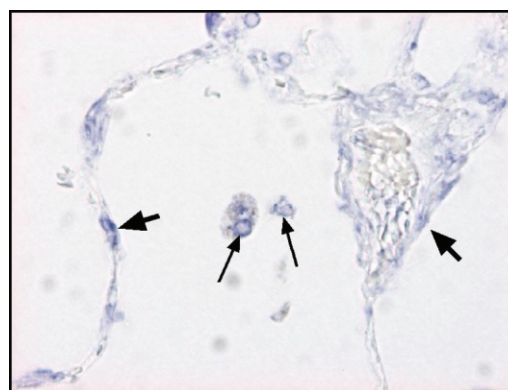
On the basis of UCSC Genome Bioinformatics, two non-overlapping genes—*ELMOD2* and *LOC152586*, encoded in the opposite directions—are located within the shared segment (fig. 2). None of the polymorphisms, identified by sequencing the susceptibility haplotype, are located on the coding regions of either one. The *ELMOD2* gene consists of nine exons that encode a 293-aa protein (EntrezGene GeneID 255520).<sup>16</sup> Exons 1–3 are located within the susceptibility haplotype (fig. 2). All nine exons and exon-intron boundaries of *ELMOD2* were sequenced in one affected patient from each genome-scan family. The sequencing revealed no mutations in the coding regions or exon-intron boundaries. We studied *ELMOD2* mRNA expression using commercially available Human Multiple Tissue cDNA panels I and II (BD Biosciences) and amplicons spanning exons 2–9, and we observed that *ELMOD2* is expressed in all studied tissues and cell types, including human lung (fig. A2 [online only]). Also, both (1) the healthy (CCL-151) and IPF-derived fibroblast (CCL-134) cell lines (data not shown) and (2) the MTC cDNA panel (BD Biosciences) representing activated and resting monocyte and lymphocyte cell types expressed *ELMOD2* (data not shown).

The other gene, *LOC152586*, is poorly characterized (En-

trezGene GeneID 152586). Several overlapping ESTs can be found at UniGene, most probably encoded by a single gene. All the identified exons are located within the susceptibility haplotype. We used the IMAGE:5267198 clone (IMAGE Consortium) that contained one of the longest potential ORFs (147 aa) for in vitro translation assay, but it failed to produce any peptide (fig. A2 [online only]). With use of Human Multiple Tissue panels I and II cDNA and amplicons covering 1,143 bp of the cDNA, the expression was identified only in testis (fig. A2 [online only]).

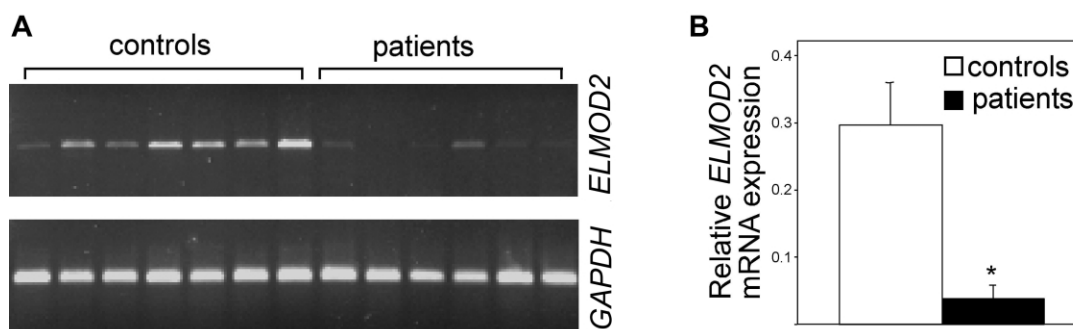
Using in situ hybridization, we recognized expression of *ELMOD2* in alveolar macrophages and alveolar walls in healthy lung (fig. 3). In contrast, even after trying several hybridization conditions, we were not able to detect any specific signal for *ELMOD2* in IPF lung samples (data not shown). We analyzed *ELMOD2* mRNA expression in IPF lung by RT-PCR on biopsies from six patients with end-stage IPF and seven healthy control individuals (fig. 4A). Snap-frozen lung biopsies were taken during surgery, and all diagnoses were verified by histology with use of the criteria presented by Katzenstein and Myers.<sup>17</sup> Individual variation occurred among both patients and controls. However, when relative expression intensities were compared, the *ELMOD2* expression level was significantly lower among the patients than among the controls ( $P = .05$ ) (fig. 4B), which supports the negative result observed by in situ hybridization.

We reported elsewhere an unusual clustering of FPF in three neighboring municipalities of Savo Province in Finland.<sup>9</sup> Five of six families informative for linkage originated from these three municipalities. This gave us a unique possibility to plan our mapping strategy on the assumption that several patients of the region, both familial and sporadic, descend from the same ancestor who lived 10–40 generations ago, which explains the geographic clustering of the disease and is similar to several other Mendelian diseases mapped successfully in Finland.<sup>12,13</sup>



**Figure 3.** By in situ hybridizations of human lung-specific mRNA, expression of *ELMOD2* found surrounding the nucleus of alveolar macrophages (*thin arrows*) and as perinuclear expression in epithelial cells on the alveolar wall (*thick arrows*).





**Figure 4.** A, *ELMOD2* mRNA expression studied with use of lung biopsies derived from patients with IPF ( $n = 6$ ) and healthy control individuals ( $n = 7$ ). B, The mean intensity of mRNA expression, significantly decreased among patients when compared with healthy controls (mean intensity of *ELMOD2*/*GAPDH* mRNA expression [ $\pm$  SEM]; the asterisk [\*] denotes  $P = .05$ ).

Our genomewide scan identified five potentially interesting regions showing either suggestive linkage or haplotype sharing among the patients. Knowing that the genome-scan families are most likely not informative enough to find linkage with genomewide significance, we settled for suggestive linkage but explored further all the genomic regions that were potentially shared by the patients. For the hierarchical fine mapping, we extended the data with the rest of familial and sporadic cases that we had identified from Savo Province. Only on chromosome 4q31.1 did we find suggestive linkage (NPL 2.1) and a conserved haplotype, which varied in size from 110 kb to 13 Mb. The shared haplotype was present in one-third of the index families, which was significantly more frequent than among the regional or population-based controls (OR 12.5, 95% CI 2.6–60.6,  $P = .0004$ ; and OR 6.3, 95% CI 2.5–15.9,  $P = .0001$ , respectively). The haplotype was found in 38% of the multiplex families and in 27% of the families in which familial background of the disease could be confirmed. In some of these seemingly sporadic patients, we could identify the shared haplotype up to 8 Mb, which suggests that at least a part of uniplex and multiplex families are related within the past 10–15 generations (fig. 2, families 17 and 21).

We identified several SNPs from the critical region, but the coding regions of the identified genes were intact among both the carriers and the noncarriers of the susceptibility haplotype. Intronic SNPs do not change the structure of the encoded protein, but they may change the expression level and splicing of genes.<sup>18,19</sup> This has been reported recently in several gene-mapping studies.<sup>20–22</sup> Although the mapping results obtained in the present study were computed in a small family cohort, the involvement of *ELMOD2* as a susceptibility gene for IPF was further supported by the decreased expression of *ELMOD2* in IPF lung compared with healthy lung. At present, the biological functions of *ELMOD2* in lung are unknown, and further studies are needed before its direct involvement in the pathogenesis of IPF can be determined. *ELMOD2* belongs to a protein family that ex-

presses a highly conserved domain found in a number of eukaryotic proteins, including Ced-12 and ELMO1-3. These molecules are known to interact in signaling pathways involved with apoptosis, phagocytosis, cell engulfment, and cell migration.

Differences in clinical manifestations and histological findings between FPF and sporadic IPF have not been studied at large, but the entities are indistinguishable.<sup>6,9,23</sup> Therefore, it is plausible to suggest that the signaling pathways present in FPF may also be of importance in sporadic IPF. Our genome scan among the Finnish families suggests a novel positional candidate. On the basis of altered expression in IPF lung and the proposed functional properties, *ELMOD2* becomes the prime candidate susceptibility gene for FPF.

### Acknowledgments

This study was supported by the Academy of Finland, Finnish Anti-Tuberculosis Association, National Technology Agency of Finland, Päivikki and Sakari Sohlberg Foundation, Sigrid Juselius Foundation, and Swedish Research Council. All participating families are gratefully acknowledged. The authors thank the staffs of the Pulmonary Clinic of Helsinki University Central Hospital, Etelä-Savo Central Hospital, and Itä-Savo Central Hospital, for helping us to find and approach the patients and their families. Dr. Marjukka Myllärniemi, Dr. Eero Sihvo, Ms. Tiina Marjomaa, and Ms. Kirsi Vuorinen are thanked for their assistance in providing us the fibroblast cell lines and the lung biopsy material. We also thank Riitta Lehtinen, Ranja Eklund, Ville-Veikko Mäkelä, and Ingegerd Fransson, for their skillful laboratory assistance.

### Web Resources

Accession numbers and URLs for data presented herein are as follows:

EntrezGene, <http://www.ncbi.nlm.nih.gov/entrez/> (for *ELMOD2* [GeneID 255520] and *LOC152586* [GeneID 152586])  
 Entrez Nucleotide, <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide&cmd=search&term=29789898> (for genomic clone NT\_016606)

Human Genome Sequence, <http://www.ncbi.nlm.nih.gov/genome/guide/human/>  
 IMAGE Consortium, <http://image.llnl.gov/> (for IMAGE clone 5267198)  
 Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for IPE, *CLGN*, and *UCPI*)  
 SISA, <http://home.clara.net/sisa/fisher.htm>  
 Staden package, <http://staden.sourceforge.net/>  
 UCSC Genome Bioinformatics, <http://www.genome.ucsc.edu/>  
 UniGene, <http://www.ncbi.nlm.nih.gov/UniGene/clust.cgi?ORG=Hs&CID=189067>

## References

- Davies HR, Richeldi L (2002) Idiopathic pulmonary fibrosis: current and future treatment options. *Am J Respir Med* 1:211–224
- Bjoraker JA, Ryu JH, Edwin MK, Myers JL, Tazelaar HD, Schroeder DR, Offord KP (1998) Prognostic significance of histopathologic subsets in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 157:199–203
- Peabody JW, Peabody JW Jr, Hayes EW (1950) Idiopathic pulmonary fibrosis: its occurrence in identical twin sisters. *Dis Chest* 18:330
- Javaheri S, Lederer DH, Pella JA, Mark GJ, Levine BW (1980) Idiopathic pulmonary fibrosis in monozygotic twins. *Chest* 78:591–594
- Uchiyama B, Matsubara N, Tominaga Y, Abe T, Satoh K, Nukiwa T (1997) Intensive interview and examination of family members revealed more than 30% incidence of familial clustering of pulmonary fibrosis. *Eur Respir J Suppl* 10:141s–142s
- Marshall RP, Puddicombe A, Cookson WOC, Laurent GJ (2000) Adult familial cryptogenic fibrosing alveolitis in the United Kingdom. *Thorax* 55:143–146
- Loyd JE (2003) Pulmonary fibrosis in families. *Am J Respir Cell Mol Biol Suppl* 3 29:S47–S50
- Steele M, Speer MC, Loyd JE, Brown KK, Herron A, Slifer SH, Burch LH, Wahidi MM, Phillips JA III, Sporn TA, McAdams HP, Schwarz MI, Schwartz DA (2005) The clinical and pathological features of familial interstitial pneumonia (FIP). *Am J Respir Crit Care Med* 172:1146–1152
- Hodgson U, Laitinen T, Tukiainen P (2002) Nationwide prevalence of sporadic and familial idiopathic pulmonary fibrosis: evidence of founder effect among multiplex families in Finland. *Thorax* 57:338–342
- Marshall RP, McAnulty RJ, Laurent GJ (1997) The pathogenesis of pulmonary fibrosis: is there a fibrosis gene? *Int J Biochem Cell Biol* 29:107–120
- Kere J (2001) Human population genetics: lessons from Finland. *Annu Rev Genomics Hum Genet* 2:103–128
- Tahvanainen E, Ranta S, Hirvasniemi A, Karila E, Leisti J, Sistonen P, Weissenbach J, Lehesjoki AE, de la Chapelle A (1994) The gene for a recessively inherited human childhood progressive epilepsy with mental retardation maps to the distal short arm of chromosome 8. *Proc Natl Acad Sci USA* 91:7267–7270
- Höglund P, Sistonen P, Norio R, Holmberg C, Dimberg A, Gustavson KH, de la Chapelle A, Kere J (1995) Fine mapping of the congenital chloride diarrhea gene by linkage disequilibrium. *Am J Hum Genet* 57:95–102
- American Thoracic Society, European Respiratory Society (2002) International multidisciplinary consensus classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med* 165: 277–304
- Kruglyak L, Daly MJ, Reeve-Daly MP, Lander E (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58:1347–1363
- Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, et al (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci USA* 99:16899–16903
- Katzenstein AL, Myers JL (1998) Idiopathic pulmonary fibrosis: clinical relevance of pathologic classification. *Am J Respir Crit Care Med* 157:1301–1315
- Pagani F, Buratti E, Stuardi C, Baralle FE (2003) Missense, nonsense, and neutral mutations define juxtaposed regulatory elements of splicing in cystic fibrosis transmembrane regulator exon 9. *J Biol Chem* 278:26580–26588
- Pagani F, Baralle FE (2004) Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet* 5: 389–396
- Karlin S, Chen C, Gentles AJ, Cleary M (2002) Association between human disease genes and overlapping gene groups and multiple amino acid runs. *Proc Natl Acad Sci USA* 99: 17008–17013
- Laitinen T, Polvi A, Rydman P, Vendelin J, Pulkkinen V, Salmikangas P, Makela S, Rehn M, Pirsanen A, Rautanen A, Zucchelli M, Gullsten H, Leino M, Alenius H, Petays T, Haahtela T, Laitinen A, Laprise C, Hudson TJ, Laitinen LA, Kere J (2004) Characterization of a common susceptibility locus for asthma-related traits. *Science* 304:300–304
- Pastinen T, Hudson TJ (2004) Cis-acting regulatory variation in the human genome. *Science* 306:647–650
- Lee H-L, Ryu JH, Wittmer MH, Hartman TE, Lymp JF, Tazelaar HD, Limper AH (2005) Familial idiopathic pulmonary fibrosis—clinical features and outcome. *Chest* 127:2034–2041
- Bitterman PB, Rennard SI, Keogh BA, Wewers MD, Adelberg S, Crystal RG (1986) Familial idiopathic pulmonary fibrosis: evidence of lung inflammation in unaffected family members. *N Engl J Med* 314: 1343–1347
- Lahiri DK, Nurnberg JI Jr (1991) A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucleic Acids Res* 19:5444